DOCUMENT RESUME

ED 306 222 TM 012 857

AUTHOR Friedland, Edward I.; Friedland, Mark W.
TITLE Common Sense about SAT Scores: The Sky Is Not Falling.
INSTITUTION Krell Software Corp., St. James, NY.
PUB DATE 88
NOTE 16p.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Viewpoints (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; Admission Criteria; *Aptitude Tests; *College Entrance Examinations; *Educational Trends; Higher Education; High Schools; Local Norms; National Norms; Predictive Measurement; *Scores; Standardized Tests; State Norms; *Test Results; Test Score Decline; Trend Analysis
IDENTIFIERS *Scholastic Aptitude Test

ABSTRACT

When properly understood, scores on the Scholastic Aptitude Test (SAT) support the hypothesis that the performance of American students as measured by the SAT is steadily improving. As the overall pool of SAT takers increases and as pressures to adopt SAT scores as an administrative measure of instructional performance mount, SAT averages at the national, state, and local levels can be expected to be driven down. Linear regression makes weeding out the impact of sample size changes on SAT averages possible. Applying these techniques to 1988 data for the 1.1 million students who took the SAT nationwide, the 913,000 who took the SAT in states where it is the principal college entrance examination, and the 187,000 who took the SAT from other states, reveals real increases in average SAT scores. From 1987 to 1988, there was a 9.1 point increase. Only one state had a decrease in real mathematics performance, and only two had decreases in real verbal performance. Many have accepted the notion of a decline in SAT scores; such a decline can be useful in competition for budgets and recognition. The apparent improvement needs to be clarified; it may be due to students being inherently better, or it may reflect that students are better prepared for taking the SAT. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED306222

Common Sense About SAT Scores

The Sky Is Not Falling

by

Edward I. Friedland, Ph.D. and Mark W. Friedland

abstract

Contrary to public belief, the slight absolute decline in SAT scores nationally does not reflect a decrease in high school student skill or preparedness. When the effects of increasing the size of the group being tested is taken into account, 1988 data show a real improvement in SAT performance.

When SAT scores are corrected to remove the effect of increased sample fraction size, they reveal a dramatic 9.1 point increase from 1987 to 1988.

A method of improving the validity of interstate comparisons of SAT scores is presented together with explanations of why scores have improved, why the improvement has not been noticed, and what must be done to make the analysis of SAT scores and trends a genuinely useful tool.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

MARK FRIEDLAND

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM012857

Ed Friedland, Ph. D, is Chairman of the Krell Software Corporation, St. James, New York. Formerly a Senior Staff Officer at the National Academy of Science and teacher at the University of California, Berkeley and State University of New York, Stony Brook, Dr. Friedland is author of The Informed Student's Guide to the SAT's, forthcoming.
Mark W. Friedland is Vice President for Corporate Planning, Krell Software.

I. The Problem

On a regular basis, data on the SAT scores of high school students are published by the College Board and the U.S Department of Education. Almost invariably, these data evoke widespread lamentation and calls for action to reverse the declining and deplorable quality of education in America.

The purpose of this paper is to show that SAT scores, properly understood, actually support the hypothesis that the performance of American students, as measured by the SAT, is steadily improving - and to explore why this is so.

II. Cream at the Top

Few people would be surprised if the average height of the tallest 5 percent of Iowa's graduating high school students were greater than the average height of the tallest 45 percent of California's graduating high school students.

Naturally, as the fraction of students drawn into the sample expands, progressively shorter students will be included, thereby decreasing the average height of the sample. Because the sample is drawn from a carefully ordered population and not at random, enlarging it adds only individuals of lower height than any already in the sample. Thus, increasing the size of the fraction sampled systematically decreases its average height.

No sensible person would bother to draw inferences about the comparative nature or nurture of those students based on the observation that students in the state where a smaller fraction was sampled proved taller, on average, than those from a state where the sample included a greater fraction of the students.

In order to make such a comparison meaningful, it would be necessary to eliminate the effect of average height attributable solely to differences in the size of the fractions sampled.

Interpreting the significance of average SAT scores on a state-by-state or district-by-district basis poses exactly the same problem.

In each state, those who take SATs are a self-selected group. As a rule, they are the cream of the academic crop - students seeking admission and/or scholarships to the best colleges. Analyzing the SAT averages of these students requires that we try to eliminate effects attributable to sample fraction size alone.

III. Percent of High School Graduates Taking the SAT as a Predictor of Average Scores

Considering the many cultural, economic, and organizational factors that shape student performance and decisions about where to apply for college entrance and the diverse pattern of educational resources available in each state, you may be amazed to discover that **more than 80% of the variation in statewide SAT averages are explained solely on the basis of variations in sample fraction size.**

As you might expect, in states where SAT's are taken by on a small fraction of high school students, average scores greatly exceed the national average. Conversely, in states where a very large fraction of high school students take SAT's, statewide average SAT scores are consistently below the national average.

This has an important policy implication. **It means that as the overall pool of SAT takers increases, and as pressures to adopt SAT scores as an administrative measure of instructional performance mount, we should expect SAT averages at the national, state, and local levels to be driven down.**

In order to truly understand the significance of SAT scores and trends, we need a way to weed out the impact of sample size changes on the SAT averages we care about. Fortunately, the simple statistical tool of linear regression lets us do just that.

IV. The Regression Lines

State-by-state data on average SAT scores and fractions of the graduating high school class taking these exams recently published by the College Board provide the information we need. Exhibit 1 shows the line that best fits the College Board data for total SAT score.

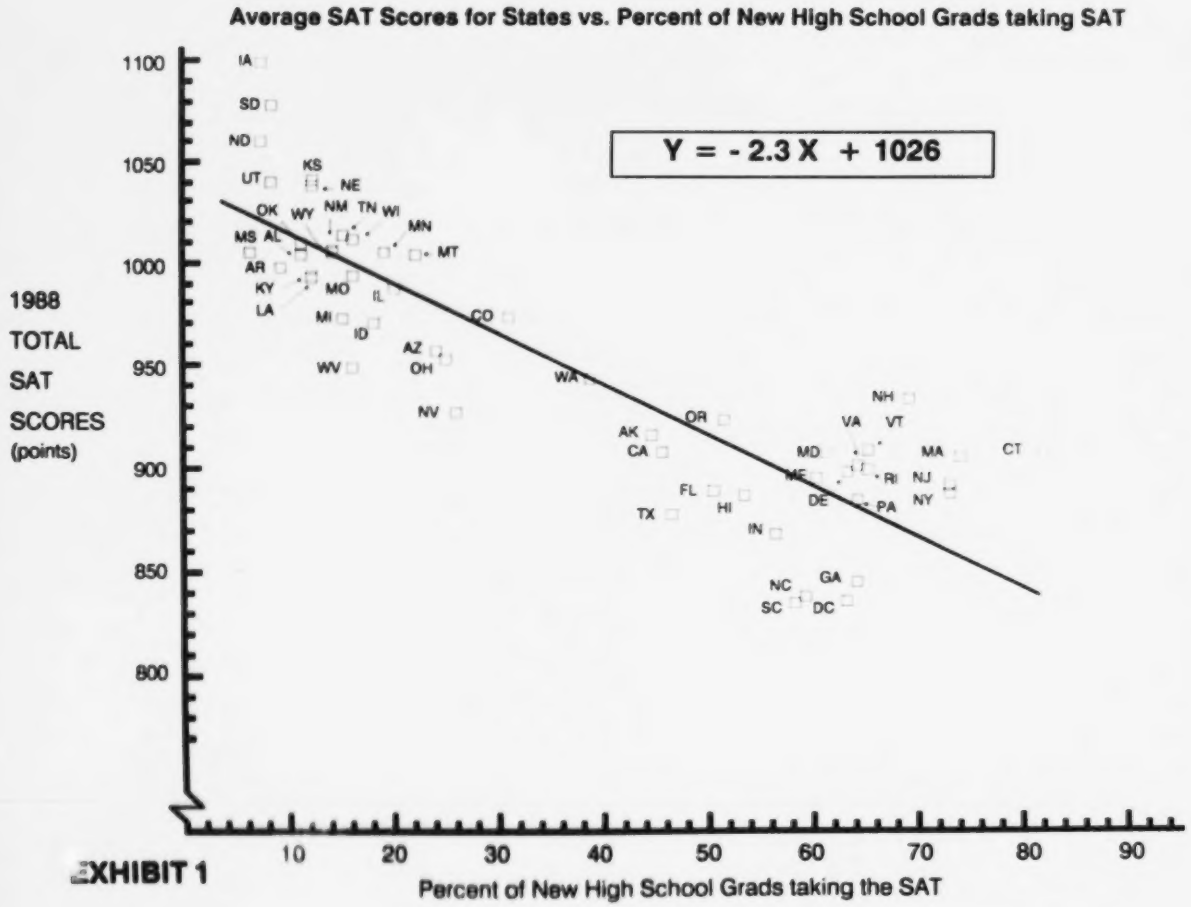

EXHIBIT 1

Total SAT score	y = 2.3 x + 1026	(r = .915)
SAT math score	y = - 1.2 x + 537	(r = .905)
SAT verbal score	y = - 1.1 x + 485	(r = .912)

Each line fits the data "best" in the sense that it minimizes the sum of the squares of the deviation of all actual 1988 SAT averages for the 50 states and District of Columbia from the SAT score that would be predicted from a knowledge of the sample fraction size alone.

These lines are computed with each data point weighted to reflect the actual number of SAT takers in each of the 51 jurisdictions.

With these lines, we can calculate the likely impact that changes in sample size are likely to have on SAT averages. For example, if the sample size in your state increases from 50 % to 60 % of the graduating seniors, you should expect an approximate decline of 23 (10 x 2.3) in average total SAT score for your state. The expected components of this decline will be (10 x 1.2) or 12 points in math and (10 x 1.1) or 11 points on the verbal portion.

V. The Measures We Present

By calculating and discounting the influence of sample fraction size changes, you can get a better grasp of the meaning of actual SAT results. For example, if the sample size in your state increases by 10 percentage points (say, from 45 to 55 percent of new high school graduates), you should expect a 23 point decrease. If the actual average decreases only 2 points, you've achieved a "real" or "corrected" increase of 21 points!

> In effect, we are defining the "real" change in an area's SAT score as:
> "Real" change = actual score - predicted score.
>
> Where, predicted score = previous score + expected change resulting from change in sample fraction size.

We believe that the two most important pieces of information needed to understand SAT scores are:

a. The "real" performance change is:

The difference between the actual score change for students in a state and the change that would be expected solely as the result of a change in the number of students taking the SAT in that state.

b. The "corrected" relative performance standing for a state is:

The difference between actual average SAT scores for students in that state and the scores that would be expected based solely on the fraction of that state's new high school graduates taking the SAT.[1] (expectation based on the nationwide 1988 data reported in Exhibit 1).

Notice that we haven't included a measure of a state's SAT averages relative to the national SAT average. While this measure provides comfort, and perhaps rhetorical advantage, in states with only a small number of SAT takers, it obscures the "true" picture of a state's performance.

The best single indicator of a state's performance is our measure, "corrected" relative performance standing. Likewise, the best measure of change and, hopefully progress, in current efforts to improve student SAT performance, is our "real" performance change.

1 See Stephan Gohmann, "Comparing SAT Scores: Problems, Biases, and Corrections," Journal of Educational Measurement, Summer, 1988. pp. 137-148. and Brian Powell and Lala Steelman, "Variations in State SAT Performance: Meaningful or Misleading?", Harvard Educational Review, Vol. 54 No. 4, November, 1984. pp. 389-412.

Table 1 presents the most recent data on these measures.

Table 1

1988 SAT Scores: Comparisons for the 22 States where SAT is the Major Exam

	1988 Current Standing			Change 1987 - 1988		Change 86-88	
State	relative	rank(51)	rank(23)	score change	rank(23)	change	rank(23)
AK	-11.4	32	14	-11	22	-1	20
CA	- 16.1	37	15	-2	17	7	16
CT	67.7	2	1	17	1	18	1
DE	15.2	17	10	-15	23	-13	23
DC	- 44.8	47	21	11	4	7	17
FL	- 23.6	42	17	0	13	6	18
GA	- 33.5	45	19	15	2	24	2
HI	- 18.7	38	16	1	12	10	13
IN	- 29.9	44	18	-2	16	9	14
ME	5.3	23	12	-2	15	13	9
MD	19.9	15	9	-3	18	11	12
MA	47.4	5	3	4	10	12	10
NH	62.9	3	2	2	11	14	7
NJ	31.4	8	4	5	8	17	5
NY	27.6	11	6	0	14	7	15
NC	- 52.0	50	22	8	6	21	3
OR	11.7	21	11	-7	21	2	20
PA	4.5	24	13	5	7	12	11
RI	20.8	13	7	9	5	14	8
SC	- 57.3	51	23	14	3	28	1
TX	- 43.7	46	20	5	9	15	6
VT	39.8	9	5	-5	20	-2	22
VA	20.5	14	8	-3	19	3	19

Sources: U.S. Department of Education, The College Entrance Examination Board, The Western Interstate Commission for Higher Education, and TIAA-CREF. Rankings, where apparently tied, are based on comparison of decimal components of score differences.

Major Findings

Our findings pertain to 3 groups of SAT takers. Group I is the 1.1 million students in the 50 states plus Washington, D.C. Group II is the 913,000 (83%) of SAT takers in the 22 states + Washington, D.C., where SAT's are the principal exam taken by college-bound high school students. Group III is the 187,000 (17%) students who took SAT's in the other 28 states.

The effect of sample fraction size is apparent in the fact that for Group III, the average SAT score was 976.5 compared to 888.7 for those in Group II. In Group III, 15.7 % percent of graduating high school students took the SAT. The 88 point score differential is readily explained by the fact that in Group II, the SAT states, 58.2 % of graduating high school students took the SAT.

Real Score Increases Outnumber Losses

Our most significant finding is that, when the effects of the expanding fraction of SAT takers among the nation's high school graduates are taken into account, **14 of the Group II states (22 SAT states + Washington, D.C.) achieved real increases in average SAT scores. In 1988, Group II increases averaged 2.7 points. The increase averages 9.1 points for Group I (the 50 states and Washington, D.C.).**

When the effects of the expanding fraction of SAT takers are taken into account, **over the 2 year span 1986 - 1988, 19 of the 22 SAT states achieved real increases in average SAT scores. The increase averaged 12.1 points for Group II (the 22 SAT states plus Washington, D.C.)**

Only One State Showed A Decrease in Real Math Performance

Of the 22 SAT states over the 2 year span, Delaware alone showed a decrease in real math performance.

In the year 1987-88, 8 of the 51 areas experienced a real decline in math scores. However, only in Delaware and Oregon were these declines more than 2 points.

Only Two States Showed Decreases in Real Verbal Performance

Of the Group I states, over the 2 year span, Delaware and Vermont showed a decrease in real verbal performance.

Improvements in Math Exceeded Improvements in Verbal Score.

In 1988, in every major SAT state, except Florida and Hawaii, math scores increased more (or decreased less) than verbal scores.

In the 51 areas, only Florida, Hawaii, Arkansas, Illinois, Kansas, Louisiana, and South Dakota showed verbal scores that increased more (or decreased less) than math scores.

We believe this reflects a combination of two factors: the impact of immigrant populations whose math performance is relatively less influenced by their relocation than is their performance on English language tests, and the greater susceptibility to improvement via teaching and coaching of SAT math skills (as opposed to SAT verbal skills).

SAT SCORE IMPROVEMENT

Presented in terms of improvement in real SAT score in the last 2 years, the major SAT states fell into these categories: (average gain = 9.1 points).

Twice average gain or more	**Better than average gain**	**Average gain or slightly less**	**Less than 1/2 average gain**
SC GA NC DC	CT ME MD MA TX RI PA NJ NH	HI CA FL NY IN	AK DE VT VA OR

Leaders, Gainers, Slippers, and States in Trouble

By taking both relative standing and real change for 1987 -1988 into account simultaneously, the states can be divided into four groups: **(SAT states are shown in bold letters.)**

Leaders - States with above average relative SAT score standings and better than average real gains in SAT score.

CT NH MA NJ PA RI - IA

Gainers - States with less than average relative SAT score standings and real gains in SAT score.

SC FL GA HI NC TX DC - AL IL NV OK

Slippers - States with above average relative SAT score standings and real losses in SAT score.

DE ME MD NY OR VT VA - CO KS MN MT NE NM ND SD TN UT WA WI WY

States in Trouble - States with below average relative SAT score standings and real losses in SAT score.

AK CA IN - AZ AR ID KY LA MI MS MO OH WV

The exact value of 1988 relative standing and real total score change are plotted in Exhibit 2.

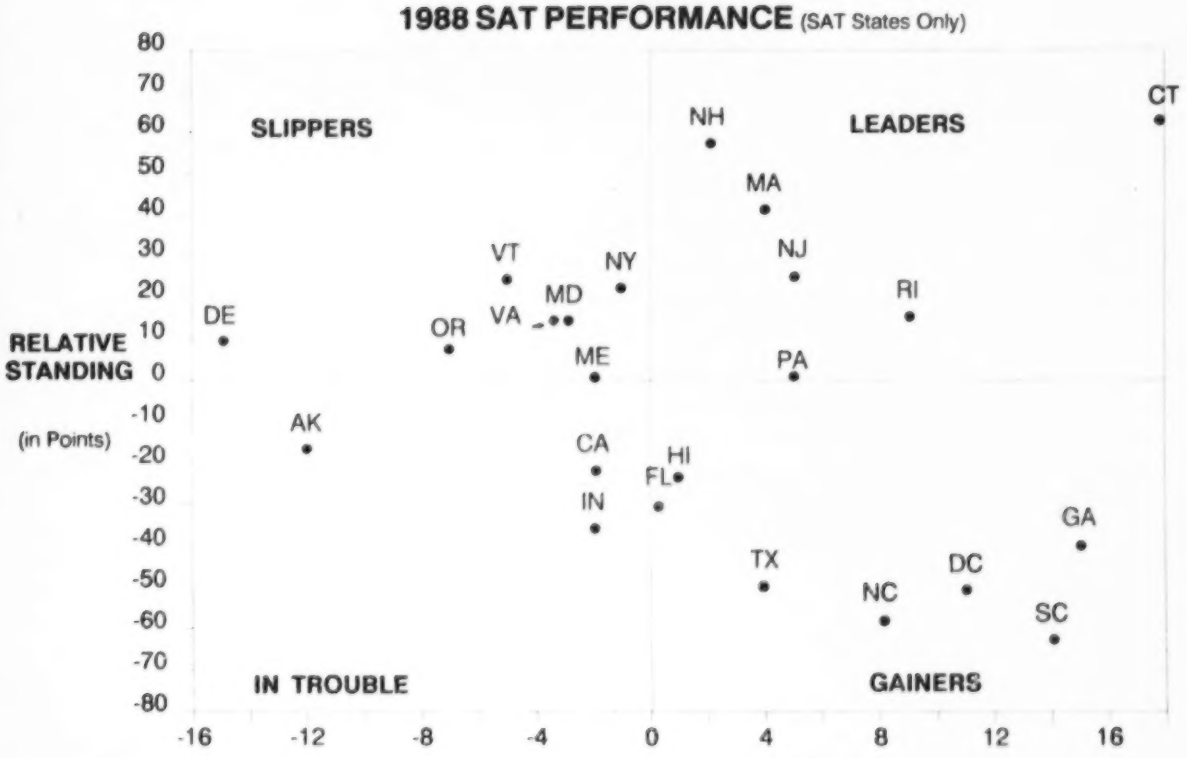

EXHIBIT 2

To put the SAT score data presented here in proper perspectice, keep in mind that each additional correct answer on the SAT contributes roughly 10 points to a student's total score. Data on both relative standing and "real" score change are expressed in "points." Thus, a state's corrected relative standing of + 20 points and a real change of 20 points represent about 2 more correct answers after correcting for sample fraction size effects.

V. So What ? The Significance of These Findings

The rather startling evidence presented here poses several major explanatory and decision problems for educators:

1. Can Meaningful Comparisons of SAT Scores Be Made ?

   (The mirror, mirror on the wall problem)

2. How to explain the increase in real score - (The Frog-prince problem)
3. How to explain the lamentation - (The Chicken Little problem)
4. How to improve learning and SAT scores - (The Goldilocks problem)

1. Comparing SAT Scores - The Mirror, Mirror on the Wall Problem

Who is the fairest of them all ?

In its most recent release of SAT performance data, the College Board includes this warning in red letters:

"MAKING COMPARISONS OF STATES OR RANKING THEM, ON THE BASIS OF SAT SCORES ALONE IS INVALID AND STRONGLY DISCOURAGED BY THE COLLEGE BOARD."

To educators this caveat has a somewhat disingenuous ring. Obviously, the purpose of standardized exams and measurements of any sort, is precisely to make comparisons possible

Indeed, once the effects of sample fraction size differences are removed, comparisons between states make good sense. The College Board's assertion that such comparisons are somehow "invalid" would be easier to accept if:

a. It were supported by reasons as to why comparisons are "invalid", and

b. The warning did not serve the economic interests of the College Board and Educational Testing Service so well. In effect, the warning reflects an attempt to counteract a most unpleasant fact of SAT marketing-namely that states are likely to resist pressing for the adoption of expanded SAT testing once they recognize that such action will likely lower their average scores.

Perhaps the SAT was not designed to facilitate comparison of the quality of education in different states, but it seems likely that SAT scores will continue to be used as one basis for comparing states as well as individual students.

As long as this is so, both individuals and educators have an obvious and coincident interest in raising SAT performance.

Similarly, everyone involved in the process of making and executing educational policy, has a parallel interest in being able to understand the ways in which SAT scores may be used to make "valid" comparisons of students and the education they receive.

While the need to correct for the influence of variations in sample fraction size in statewide comparisons is clear, another far more troubling problem exists - the need to make valid inferences about SAT score trends.

A Statistical Observation - Measurement and Purpose

SAT's have been criticized in many ways. Yet, the bottom line is that SAT's endure because college admissions officers find SAT scores provide one useful basis for explaining and rationalizing admissions decisions, not because the exams themselves or the ways in which the SAT's are used are free of major defects.

The simple truth is that the testing technology which could rapidly and inexpensively provide the most detailed and subtle measures of a student's knowledge and mental skills already exists. However, changing the SAT system to provide such data to colleges would seem to offer no advantage to colleges insofar as improving their ability to provide post facto explanations of their admissions decisions.

In naked terms, the Educational Testing Service derives its SAT revenues only because colleges continue to require applicants to take the exam. Whatever the needs and desires of other educators may be for information about the performance of students and/or institutions, the ETS is likely to accord priority to the purposes of the college administrators from whom their blessings ultimately flow.

An Elastic Tapemeasure ?

How serious the defects that exist in the nature of the SAT's or the SAT system itself, may be debatable, but our need for reliable tools to measure student performance is not. **We cannot begin to claim a true understanding of the significance of any trends in SAT scores until we have some confidence in the reliability of the SAT as a measuring device.**

The reliability issue has nothing to do with how valuable we believe the peculiar combination of math and English language skills SAT's measure may be, it is based on the conviction that **we have a right to expect some way of accurately comparing the difficulty of SAT's administered at different times.**

This problem is similar to that faced by economists trying to analyze trends. Where time series information is expressed in dollars, in order to be meaningful, comparisons must take inflation in account.

With SAT scores, we need a way of correcting for score "inflation." Just as the value of a dollar has changed over time, so has the difficulty of SAT examinations.

Conceptually, it's quite an easy task to measure exactly how the difficulty of SAT exams have varied over the years. **Doing so requires nothing more than administering past and present SAT's to a single control group.**

Regrettably, if such tests have been conducted, no published information on their results exists. Unquestionably, we need a way to "calibrate" SAT scores by accurately compensating for variations in exam difficulty over time. Now is the time to conduct the simple study and obtain the 'difficulty' data needed by everyone who wants to understand SAT score trends.

In a nutshell, even after the the effect of sample size has been removed, we still won't know how to interpret SAT score trends until we can accurately measure and then eliminate whatever "inflation" exists in the SAT exam itself.

2. Why the Increase in Real SAT Scores ? - The Frog Prince Problem

Since average SAT scores show a real increase (after the effects of sample fraction size change are eliminated), we face what may be called "the frog-prince problem." Namely, how to explain the very substantial increase in SAT scores. If so many students (frogs) are revealed by their SAT scores to actually be princes, what is causing this improvement. Let's examine 2 major possibilities that may be true individually or in combination.

A. Maybe Students Are Inherently Better

This is always a possibility. Unfortunately, we just don't have the evidence we'd need to prove or disprove hypotheses about students being "better" in any relevant sense. Whether students are smarter, more motivated, or "better" in any other fundamental way cannot be determined.

One approach to explaining SAT score trends has been to argue that the main source of change has been a change in 'who' takes the exam. Our argument that sample fraction size is critical falls into this category. So do the variety of explanations regarding changes in the composition of the SAT-taking group offered by the College Board in **On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline** (New York, College Board, 1977).

The unstated message that accompanies this approach is "educators and schools should not be blamed for SAT score declines." Going one step further is Harold Howe II's insightful article, "Let's Have Another SAT Score Decline," (**Phi Delta Kappan,** May, 1985.) Howe correctly suggests the declines in SAT averages should be recognized as a reflection of greater inclusiveness. Simply put, we should be pleased to realize that the decline of uncorrected SAT scores actually signifies a growth of student aspirations that translate into the expanded sample fraction size.

B. Maybe Students Are Better Prepared for the SAT

The rise in SAT scores corrected for sample fraction size changes can be explained in a very different way. We can argue that by correcting for sample fraction size, we effectively cast out the influence of changes in the composition of the test-taking group. Then, by studying the corrected SAT score changes, we can determine how well students are prepared.

But the term "prepared" has many senses. To some "prepared" comprehends the entire education process. Viewed in this light, SAT score are often seen as a general measure of instructional performance. While we would all love to have such a measure, SAT scores just won't do.

Virtually no one would be willing to accept the idea that the quality of instructional performance in general can be measured by how well students have mastered the odd mixture of mathematics and English language agility which the SAT actually tests. Moreover, it would even be a mistake to presume SAT scores can tell us how well our regular educational institutions prepare students in those skills tested by the SAT. Indeed, the mistake is obvious when we look at the far more limited idea of "prepared" as meaning "prepared to take the SAT." Here are some facts to consider:

1. During the last 4 decades a large private SAT preparation industry has developed by meeting the demand for SAT coaching.
2. Almost one half the high schools surveyed by the College Board reported offering courses to prepare students for the SAT.[2]
3. The ETS, Educational Testing Service, creator and administrator of the SAT now grudgingly acknowledges that coaching can improve SAT scores.
4. The use of computer software to effectively prepare students for the SAT is widespread.[3]
5. New economic arrangements, such as software licensing, foster the wider availability of SAT preparation tools.
6. 1 in 9 students report taking private SAT prep courses. A majority of those indicated their belief that they were "helped a lot."[2]

Some argue against such SAT coaching citing their beliefs that it is inherently inequitable because of its cost, that somehow it "distorts" the true picture of student skill, or that, still worse, it diverts teaching effort away from fundamental issues and toward exam performance trivia.

Virtually everyone, however, recognizes that, like athletic coaching and practice, SAT preparation helps individuals raise their SAT scores. It seems clear that SAT scores (corrected for sample fraction size and inflation effects) will continue to rise as access to SAT exam preparation widens.

3. Lamentations - The Chicken Little Problem

How could educators and the general public overlook so obvious and dominant a factor as the effect of increasing sample size ?

The answer is, because it suits their purposes. No cynicism is intended here, just the recognition of a simple truth. Denouncing the decline of education has been a fashionable pastime throughout human history.

In our democracy, partisans of various causes clamor for public attention to parcular"problems." Indeed, the first step to success for any "cause" in America is to promote public awareness of the very existence of circumstances defined as a "problem." Ultimately, attention and funding are

2 "1 in 9 Said to Use Private Coaching for SAT", Education Week, October 12, 1988, p.6.

3 Krell's SAT Prep package, for example, provides personal tutoring based on continuous analysis of the individual student's performance on practice problems. Sold to individuals with a money-back guarantee if less than an 80 point SAT score increase was attained, Krell's software had less than a 1% refund rate.

directed to those causes whose advocates are most successful at persuading others of the importance and potential menace of their 'problem.'

Little wonder then, that the public is bombarded with endless variations on the theme "the sky is falling!"

A common truth shared by political activists of every conviction and varying degrees of sincerity, is that the best way to advance a cause is by convincing others of the dire threat posed by a particular "problem."

Educators, like those in the law-enforcement, military, medical, and other professional communities, have all learned the same lesson: **financial and professional rewards go, not to those people and organizations that have succeeded in dealing with problems, but to those whose problems have grown more menacing.**

In this light, it seems easy enough to understand how readily many are willing accept as "true" the notion of a decline in SAT scores. Such a "truth" can be quite useful in the battle for education budgets and recognition.

4. How to Improve Learning and Scores - The Goldilocks Problem

When we said earlier that no cynicism was intended in explaining the pervasive failure to "see" the gain in "real" SAT scores, we weren't being facetious. For most of us, and especially for educators, the primary goal is not to describe reality with exactitude, but to change it for the better.

This observation is not meant to justify some platonic "deception in the cause of education." Certainly, it never can justify self-deception about the realities of education and our need to know them.

Improving learning and SAT scores are complementary objectives - and can be achieved only if we secure the necessary resources and invest them wisely. This may suggest an undue focus on economic considerations. However, whether or not we find the idea felicitous, the world is arranged so that failure to attend to economic facts, like failure to attend to the force of gravity, is a sure prescription for disaster.

We must confront the need to provide both education in its most profound sense and those mundane and tangible rewards to students and parents - rewards such as SAT scores which translate into expanded opportunities in life.

This poses the "Goldilocks Problem." That is, the problem of getting the mix of our educational efforts "just right."

Fortunately for all of us, new economic facts make it possible to provide the benefits of higher scores on SAT's and other standardized tests at less expense than ever. Parents, students, and funding agencies will all be delighted to discover that resources can be shifted toward the most essential areas of instruction because we can now cheaply expand access to effective SAT preparation.

For example, our firm, Krell Software, offers school districts the right to reproduce our best personalized tutorial software for SAT, ACT, and other standardized exam preparation for about one dollar per student. License agreements encourage the widest possible dissemination of these exam preparation resources including personal copies for home use.

To be sure, interpreting and raising SAT scores are but one tiny facet of the overall task of education. But both can and must be done much better than they are presently.

Some very apparent gaps in our understanding of what SAT scores "really mean" require prompt remedies. This is the time to carry out the simple tests to make analysis of SAT trends possible by revealing whatever "inflation" exists in the exams themselves. It is also the time to demonstrate that valid state-to-state score comparisons can be made by correcting for sample fraction size effects.

The public and professionals in the education community know that SAT scores cannot accurately reveal which individuals or states are "fairest of them all." But, instead of providing warning labels, the College Board or some independent organization should publish the data described here - the data necessary to make valid state-to-state comparisons and analyses of trends.